A New Framework to Quantify the Uncertainty in Inverse Problems

Wenlong Zhang

Department of Mathematics, Southern University of Science and Technology (SUSTech)

Joint work with Zhiming Chen and Jun Zou

March 25 2023

Observational data

$$y_i = (Fu)(x_i) + e_i$$

where e_i is random variable. It's to recover u from y.

For industrial consideration:

- Data based observation: point-wise or patch measurement.
- Uncertain measurements: some amount of noisy data.

Take $F = I$ for example: imaging, denoising, surface fitting.

F to be forward operator, it's inverse problem.

Image processing

$$y_i = u(x_i) + e_i$$

- Wavelets, Rudin-Osher-Fatemi (ROF) model
- Compressed sensing(sparse method), e.g. Terence Tao
- Neural Networks

The observational data

The observational data: $y_i = u_0(x_i) + e_i,\ 1 \le i \le n$. u_0 comes from partial differential equations.

Given:

- noise e_i
- Large n, e.g. $n = 10^6$

Question:

- Recover u_0
- Error estimate

Applications: Data mining, interpolation, surface fitting, ...

Thin plate spline model

Data $y_i = u_0(x_i) + e_i,\ 1 \le i \le n$, Ω bounded domain of $\mathbf{R}^d$, $d \le 3$.

D^2-spline:

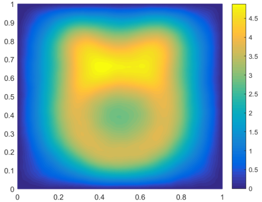
$$\min_{u \in H^2(\Omega)} \frac{1}{n} \sum_{i=1}^{n} (u(x_i) - y_i)^2 + \lambda_n |u|^2_{H^2(\Omega)},$$

where $\lambda_n > 0$.

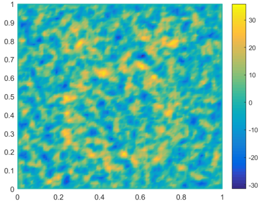
- The choice of λ_n
- Discrete method
- Error estimate

The choice of λ_n

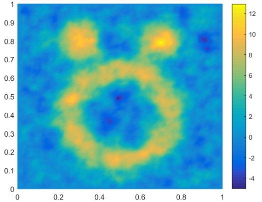
Take inverse parabolic source problem for example:



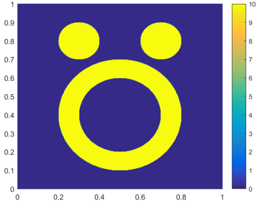
(a) $\lambda_n = 10^{-4}$ too big, over smooth



(b) $\lambda_n = 10^{-7}$ too small, over fit



(c) $\lambda_n = 1.5 \times 10^{-6}$



(d) true image

Error estimate

If $\{e_i\}_{i=1}^n$ s.t. $\mathbb{E}[e_i] = 0$ and $\mathbb{E}[e_i^2] \leq \sigma^2$. Define $\|u\|_n^2 = \frac{1}{n}\sum_{i=1}^n u^2(x_i)$.

Theorem

Let $u_n \in H^2(\Omega)$ be the unique solution of the thin plate spline model. Then there exist constants $\lambda_0 > 0$ and $C > 0$ s.t. for any $\lambda_n \leq \lambda_0$ and $n\lambda_n^{d/4} \geq 1$,

$$\mathbb{E}\left[\|u_n - u_0\|_n^2\right] \leq C\lambda_n |u_0|_{H^2(\Omega)}^2 + \frac{C\sigma^2}{n\lambda_n^{d/4}},$$

Optimal smoothing parameter:

$$\lambda_n^{1+d/4} = O((\sigma^2 n^{-1})|u_0|_{H^2(\Omega)}^{-2}).$$

Finite element for spline model

e.g. $u_0 = (xy)^{1.501} \in H^2((0,1)\times(0,1))$, $n = 10^6$, $\sigma = 0.1$.

$$\lambda_n^{1+d/4} = O((\sigma^2 n^{-1})|u_0|_{H^2(\Omega)}^{-2}).$$

Optimal $\lambda_n \approx 2 \times 10^{-6}$,
Mesh size $h = O(\lambda_n^{1/4}) \approx 0.04$

$$Ax = y$$

- Radial basis: A, $10^6 \times 10^6$, full matrix
- **Finite element method:** A, 3000×3000, sparse matrix

Numerical examples

Optimal λ_n:

$$\lambda_n^{1/2+d/8} = O(\sigma n^{-1/2}(|u_0|_{H^2(\Omega)})^{-1}),$$

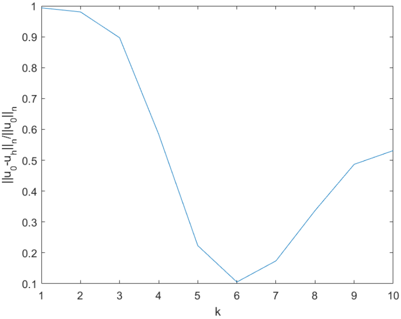


Figure: $u_0 = \sin(2\pi x^2 + 3\pi y)e^{x^3+y}$. *The relative error* $\|u_h - u_0\|_n / \|u_0\|_n$ *for different* $\lambda_n = 10^{-k}$. $\sigma n^{-1/2} = 1/50$. $\lambda_n^{opt} \approx 2.4 \times 10^{-6}$.

The smoothing parameter λ_n

$$\lambda_n^{1/2+d/8} = O(\sigma n^{-1/2}(|u_0|_{H^2(\Omega)})^{-1}).$$

- σ, u_0 unknow
- How to determine λ_n

Algorithm

(Self-consistent algorithm for finding λ_n)

1° *Initial guess* $\lambda_{n,0}$;

2° *For* $k \geq 0$ *and* $\lambda_{n,k}$, *compute* u_h *with* $\lambda_{n,k}$ *and* $h = \lambda_{n,k}^{1/4}$;

3° *Update* $\lambda_{n,k+1}^{1/2+d/8} = \|u_h - y\|_n n^{-1/2}(|u_h|_{2,h})^{-1}$.

The smoothing parameter λ_n

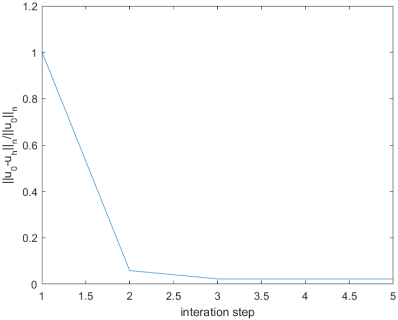


Figure: $u_0 = \sin(2\pi x^2 + 3\pi y)e^{x^3+y}$, $n = 10^6$, $\sigma = 1$. $\lambda_{n,5} = 4.3496e - 08$, *the optimal* $\lambda_n = 4.5054e - 08$

General model

$$m_i = (Sf^*)(x_i) + e_i, \;\; i = 1, 2, \cdots, n,$$

where $e = (e_1, e_2, \cdots, e_n)^T$ is the data noise vector, with $\{e_i\}_{i=1}^n$ being independent random variables. S is the forward operator from X to Y.

We look for an approximate solution f_n of the unknown source function f^* through the least-squares regularized minimization:

$$\min_{f \in X} \frac{1}{n} \sum_{i=1}^{n} |(Sf)(x_i) - m_i|^2 + \lambda_n \|f\|_X^2,$$

where $\lambda_n > 0$ is called a regularization parameter.

General model

Assumption

We assume that

(1) There exists a constant $\beta > 1$ such that for all $u \in Y$,

$$\|u\|^2_{L^2(\Omega)} \le C(\|u\|^2_n + n^{-\beta}\|u\|^2_Y), \quad \|u\|^2_n \le C(\|u\|^2_{L^2(\Omega)} + n^{-\beta}\|u\|^2_Y). \quad (1)$$

(2) The first n eigenvalues, $0 < \eta_1 \le \eta_2 \le \cdots \le \eta_n$, of the eigenvalue problem

$$(\psi, v)_X = \eta\,(S\psi, Sv) \ \ \forall v \in X,$$

satisfy that $\eta_k \ge Ck^\alpha$ $(k = 1, 2, \cdots, n)$ for some constant C depending only on the operator $S : X \to Y$ and the index α such that $1 < \alpha \le \beta$.

General model

Theorem

Let $f_n \in X$ be the unique solution of the inverse problem. Then there exist constants $\lambda_0 > 0$ and $C > 0$ such that for any $\lambda_n \leq \lambda_0$,

$$\mathbb{E}\left[\|Sf_n - Sf^*\|_n^2\right] \leq C\lambda_n \|f^*\|_X^2 + C\sigma^2/(n\lambda_n^{1/\alpha}),$$

$$\mathbb{E}\left[\|f_n\|_X^2\right] \leq C\|f^*\|_X^2 + C\sigma^2/(n\lambda_n^{1+1/\alpha}).$$

General model

Assumption

For a unit ball SY in Y and any $\varepsilon > 0$, there exists a constant $\gamma < 2$ such that the covering entropy is controlled by

$$\log N(\varepsilon, SY, \|\cdot\|_{L^\infty(\Omega)}) \leq C\varepsilon^{-\gamma}.$$

General model

Theorem

Let $\rho_0 = \|f^*\|_X + \sigma n^{-1/2}$*, and* $f_n \in X$ *be the solution of the minimization problem. If we take* $\lambda_n^{1/2+\gamma/4} = O(\sigma n^{-1/2} \rho_0^{-1})$*, then there exists a constant* $C > 0$ *such that*

$$\mathbb{P}(\|Sf_n - Sf^*\|_n \geq \lambda_n^{1/2} \rho_0 z) \leq 2\, e^{-Cz^2} \text{ and } \mathbb{P}(\|f_n\|_X \geq \rho_0 z) \leq 2\, e^{-Cz^2}.$$

General model

We can directly verify that the solution $f_n \in X$ satisfies the weak formulation

$$\lambda_n(f_n, v)_X + (Sf_n, Sv)_n = (m, Sv)_n \quad \forall v \in X\,.$$

Let $V_h \subset X$ and $Y_h \subset Y$ be two discrete function spaces (e.g., finite element spaces) with dimensions N_h and M_h.

$S_h : X \to Y_h \subset Y$ be the discrete approximation.

General model

Assumption

For the discrete operator $S_h : X \to Y_h \subset Y$,
(1) there exists an error estimate $e(h)$ *such that the discrete operator* S_h *satisfies*

$$\|Sf - S_h f\|_n^2 \le Ce(h)\|f\|_X^2 \;\forall f \in X\,.$$

(2) For any $f \in X$, *there exists* $v_h \in V_h$ *such that*

$$\lambda_n\|f - v_h\|_X^2 + \|S_h f - S_h v_h\|_n^2 \le C(\lambda_n + e(h))\|f\|_X^2.$$

General model

Theorem

$$
\begin{aligned}
\mathbb{E}\left[\|Sf^* - S_h f_h\|_n^2\right] &\leq C(\lambda_n + e(h))\|f^*\|_X^2 \\
&\quad + C\left[1 + \frac{e(h)}{\lambda_n} + \frac{N_h e(h)}{\lambda_n^{1-1/\alpha}}\right]\frac{\sigma^2}{n\lambda_n^{1/\alpha}}, \\
\mathbb{E}\left[\|f^* - f_h\|_X^2\right] &\leq C\frac{\lambda_n + e(h)}{\lambda_n}\|f^*\|_X^2 \\
&\quad + C\left[1 + \frac{e(h)}{\lambda_n} + \frac{N_h e(h)}{\lambda_n^{1-1/\alpha}}\right]\frac{\sigma^2}{n\lambda_n^{1+1/\alpha}}.
\end{aligned}
$$

In particular, if $e(h) \leq C\lambda_n$ *and* $N_h e(h) \leq C\lambda_n^{1-1/\alpha}$*, we have*

$$
\begin{aligned}
\mathbb{E}\left[\|Sf^* - S_h f_h\|_n^2\right] &\leq C\lambda_n\|f^*\|_X^2 + C\sigma^2/(n\lambda_n^{1/\alpha}), \\
\mathbb{E}\left[\|f^* - f_h\|_X^2\right] &\leq C\|f^*\|_X^2 + C\sigma^2/(n\lambda_n^{1+1/\alpha}).
\end{aligned}
$$

General model

Theorem

Let $f_h \in V_h$ be the solution of discrete problem. Denote by $\rho_0 = \|f^*\|_X + \sigma n^{-1/2}$. If we take $e(h) \le C\lambda_n$, $N_h e(h) \le C\lambda_n^{1-\gamma/2}$ and $\lambda_n^{1/2+\gamma/4} = O(\sigma n^{-1/2}\rho_0^{-1})$, then there exists a constant $C > 0$ such that for any $z > 0$,

$$\mathbb{P}(\|S_h f_h - Sf^*\|_n \ge \lambda_n^{1/2}\rho_0 z) \le 2e^{-Cz^2} \text{ and } \mathbb{P}(\|f_h\|_X \ge \rho_0 z) \le 2e^{-Cz^2}.$$

Inverse source problem in $\mathbb{R}^d$, $d \leq 3$

$$\begin{cases} u_t + Lu = f(x)g(t) \quad \text{in } \Omega \times (0, T), \\ u(x,t) = 0 \quad \text{on } \partial\Omega \times (0, T), \qquad u(x,0) = 0 \quad \text{in } \Omega, \end{cases}$$

Here $Lu = -\nabla \cdot (a(x)\nabla u) + c(x)u$ and $Sf = u(\cdot, T)$ is final time measurements.

The inverse problem is to recover the source $f(x)$ form the final time measurements:

$$y_i = Sf(x_i) + e_i$$

Inverse source problem

The least-squares regularized minimization:

$$\min_{f \in L^2(\Omega)} \frac{1}{n} \sum_{i=1}^{n} |(Sf)(x_i) - m_i|^2 + \lambda_n \|f\|_{L^2(\Omega)}^2,$$

1. Eigenvalue distributions of elliptic operators —— Expectation

$$L\psi = \mu\,\psi \;\text{ in } \Omega, \quad \psi = 0 \;\text{ on } \partial\Omega$$

   has a countable set of positive eigenvalues $C_1 k^{2/d} \leq \mu_k \leq C_2 k^{2/d}$.

2. Covering number of function space——Exponential decay tail

$$\log N(\varepsilon, SW^{2,2}(Q), \|\cdot\|_{L^\infty(Q)}) \leq C\varepsilon^{-1},$$

Inverse source problem

Theorem

For the minimizer $f_n \in L^2(\Omega)$, there exist constants $\lambda_0 > 0$ and $C > 0$ such that the following estimates hold for any $\lambda_n \leq \lambda_0$:

$$\begin{aligned} \mathbb{E}\left[\|Sf_n - Sf^*\|_n^2\right] &\leq C\lambda_n \|f^*\|^2_{L^2(\Omega)} + C\sigma^2/(n\lambda_n^{d/4}), \\ \mathbb{E}\left[\|f_n\|^2_{L^2(\Omega)}\right] &\leq C\|f^*\|^2_{L^2(\Omega)} + C\sigma^2/(n\lambda_n^{1+d/4}). \end{aligned}$$

Moreover, if $\lambda_n \geq n^{-4/d}$ and $g > 0$ in $[0, T]$, then

$$\mathbb{E}\left[\|f_n - f^*\|^2_{H^{-1}(\Omega)}\right] \leq C\lambda_n^{1/2} \|f^*\|^2_{L^2(\Omega)} + C\sigma^2/(n\lambda_n^{1/2+d/4}).$$

Inverse source problem

Theorem

Let $\rho_0 = \|f^*\|_{L^2(\Omega)} + \sigma n^{-1/2}$*. If we take* λ_n *such that* $\lambda_n^{1/2+d/8} = O(\sigma n^{-1/2}\rho_0^{-1})$*, then the following estimates hold for some constant* $C > 0$*:*

$$\mathbb{P}(\|Sf_n - Sf^*\|_n \geq \lambda_n^{1/2}\rho_0 z) \leq 2e^{-Cz^2}, \quad \mathbb{P}(\|f_n\|_{L^2(\Omega)} \geq \rho_0 z) \leq 2e^{-Cz^2}.$$

Moreover, if $\lambda_n \geq n^{-4/d}$ *and* $g > 0$ *in* $[0, T]$*, then*

$$\mathbb{P}(\|f_n - f^*\|_{H^{-1}(\Omega)} \geq \lambda_n^{1/4}\rho_0 z) \leq 2e^{-Cz^2}.$$

Finite element method

We use the backward Euler scheme

$$\Big(\frac{u_h^i - u_h^{i-1}}{\tau}, v_h\Big) + a(u_h^i, v_h) = (fg^i, v_h) \quad \forall v_h \in V_h,$$

where $a(v, w) = (a\nabla v, \nabla w) + (cv, w)$ for any $v, w \in H_0^1(\Omega)$.
The classical theory requires the regularity $\partial_{tt} u \in L^1(0, T; L^2(\Omega))$ of the solution of the problem, but this will not be guaranteed in this case. We show that

$$\|S_{\tau,h} f - Sf\|_{L^2(\Omega)} \leq C(h^2 + \tau|\ln\tau|)\|f\|_{L^2(\Omega)},$$

Finite element method

Theorem

Let $g \in H^2(0,T)$. $\{e_i\}_{i=1}^n$ are independent random variables satisfying $\mathbb{E}[e_i] = 0$ and $\mathbb{E}[e_i^2] \le \sigma^2$. Then there exist constants $\lambda_0 > 0$ and $C > 0$ such that for any $\lambda_n \le \lambda_0$ and $\tau|\ln\tau| = O(h^2)$, the following estimates hold:

$$\mathbb{E}\big[\|Sf^* - S_{\tau,h}f_h\|_n^2\big] \le C(\lambda_n + h^4)\|f^*\|_{L^2(\Omega)}^2 + C\left(1 + \frac{h^4}{\lambda_n}\right)\frac{\sigma^2}{n\lambda_n^{d/4}}.$$

Moreover, if $\lambda_n \ge n^{-4/d}$ and $g > 0$ in $[0,T]$, we have

$$\mathbb{E}\big[\|f^* - f_h\|_{H^{-1}(\Omega)}^2\big] \le C(\lambda_n^{1/2} + h^2)\Big(1 + \frac{h^4}{\lambda_n}\Big)\|f^*\|_{L^2(\Omega)}^2 + C(\lambda_n^{1/2} + h^2)\Big(1 + \frac{h^4}{\lambda_n}\Big)\frac{\sigma^2}{n\lambda_n^{1+d/4}}.$$

Optimal choice

Optimal parameter choice predicted by theory in $\mathbb{R}^2$:

$$\lambda_n^{3/4} = \sigma n^{-1/2} \|f^*\|_{L^2(\Omega)}^{-1}.$$

Algorithm (Computing an estimate of the regularization parameter λ_n)

1° *Given an initial guess of* $\lambda_{n,0}$*; for* $j = 0, 1, \cdots,$ *do the following*
2° *Solve regularization problem for* f_h *with* λ_n *replaced by* $\lambda_{n,j}$ *over the mesh* $\mathcal{M}_h$*;*
3° *Update* $\lambda_{n,j+1}$*:* $\quad \lambda_{n,j+1}^{1/2+d/8} = n^{-1/2} \|S_{\tau,h} f_h - m\|_n \, \|f_h\|_{L^2(\Omega)}^{-1}.$

Numerical examples

We will test on the following L^2 function with no more derivatives, since we do not assume any further source condition:

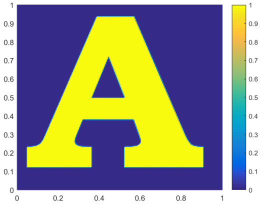


Figure: *The surface plot of the exact solution f^*.*

Numerical examples

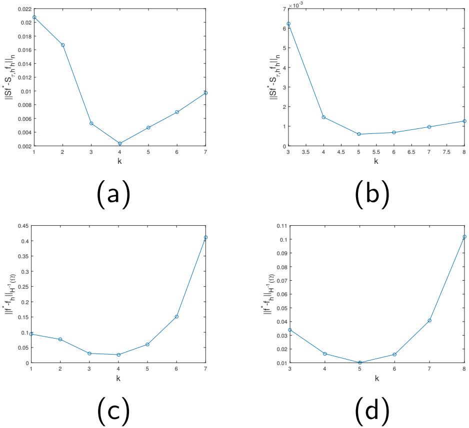
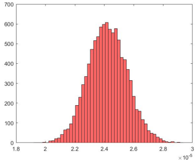
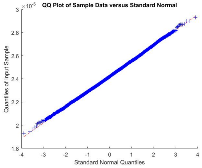


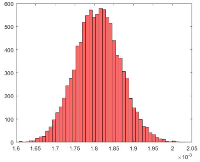
Figure: *Optimal choice are* $\lambda_n \approx 2.3 \times 10^{-4}$ *(for* $\sigma = 0.1$ *(a) and (c)) and* $\lambda_n \approx 1.1 \times 10^{-5}$ *(for* $\sigma = 0.01$ *(b) and (d)).*



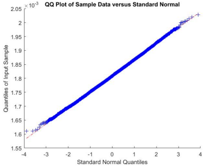
(a)



(b)



(c)



(d)

Figure: *(a) and (b) are the histogram (left) and quantile-quantile (right) plots of the empirical error* $\|S_{\tau,h}f_h - Sf^*\|_n$ *with* $10,000$ *samples. (c) and (d) are the histogram (left) and quantile-quantile (right) plots of the error* $\|f_h - f^*\|_{H^{-1}(\Omega)}$ *with* $10,000$ *samples.*

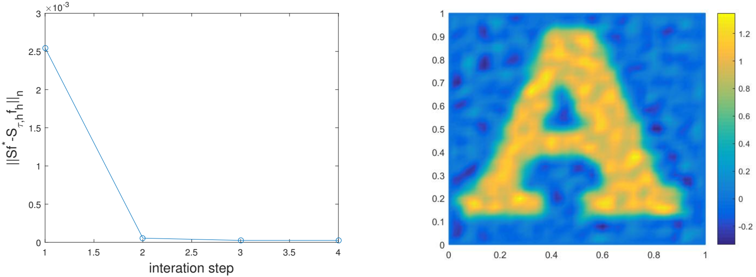
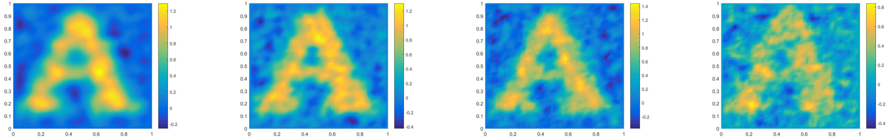


Figure: *The relative empirical error* $\|Sf^* - S_{\tau,h}f_h\|_n$ *at each iteration (left); The computed solution* f_h *at the end of iterations (right).*

Numerical examples



(a) (b) (c) (d)

Figure: *(a)-(d) are the computed solutions f_h when $T = 1,\ 0.1\ , 0.01,\ 0.001$, respectively.*

References

1. Z. Chen, R. Tuo and W. Zhang, Stochastic Convergence of A Nonconforming Finite Element Method for the Thin Plate Spline Smoother for Observational Data, SIAM Journal on Numerical Analysis, 2018, 56: 635-659.
2. Zhiming Chen, Wenlong Zhang, Jun Zou, Stochastic convergence of regularized solutions and their finite element approximations to inverse source problems, SIAM Journal on Numerical Analysis, 2022, 60(2), 751-780.

Thank you